THE PSYCHOLOGICAL STUDY OF GENDER: MOVING BEYOND THE BINARY

ALEXIS J. STRAZDS

A THESIS SUBMITTED TO
THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF ARTS

GRADUATE PROGRAM IN PSYCHOLOGY
YORK UNIVERSITY
TORONTO, ONTARIO

DECEMBER 2018

Abstract

Quantitative gender research within Psychology has traditionally been limited and inherently exclusionary of people with non-binary genders. Focusing research on aspects of gender with inclusive methods can allow us to develop a more fundamental understanding of gender beyond just masculinity and femininity. The present study explores the validity of two measures of gender self-schema – gender salience and schematic breadth – to demonstrate how such methods can be both more inclusive and lead to a more nuanced understanding of gender. Schematic breadth was explored separately between three types – role, presentation/performance, and identity – and between gender groups. Notably, non-binary people were found to use significantly fewer role terms than other genders. This suggests that gender role may take on a different meaning for non-binary people, supporting the idea that new approaches are required to understand gender in an inclusive way that moves beyond the binary of masculinity and femininity.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

The Psychological Study of Gender: Moving Beyond the Binary

Gender is a nuanced concept, and countless genders have arisen that challenge the traditional view of gender as a strict, immutable, and biologically-based binary (Butler, 1990, 1993; Corwin, 2009; Fausto-Sterling, 2000; Hansburry, 2005). The experiences of transgender people, whose gender identity differs from that assigned at birth, demonstrates that gender is not immutable (e.g., Bornstein, 1992). Gender Fluid individuals, those who move between gender identities or presentations, demonstrate that gender is not only mutable, but can be constantly changing for an individual. Furthermore, some transgender (often shortened to trans) people identify as something outside of the binary of man/masculine/male and woman/feminine/female, claiming identities that are either somewhere in between, a combination of both, neither, or something else entirely. For some, gender identity is political, and they may claim identities like Genderqueer or Genderfuck.[1] While researchers have begun to challenge the gender binary, approaches to the quantitative study of gender have previously operated on a very narrow and binary conceptualization of gender, which cannot adequately account for the many genders described above (Hyde, Bigler, Tate, & Anders, 2018; Moradi & Parent, 2013; Westbrook & Saperstein, 2015). In an effort to begin to address this problem, this study will explore the limitations of previous approaches to the study of gender, and propose, develop, and validate two measures of aspects of gender that address these limitations and are more inclusive of all genders.

---

[1] It should be noted that these terms as used here are rooted in contemporary Western English-speaking cultures, and this work does not claim to speak to genders of other cultures, languages, or times.

Previous Approaches to the Study of Gender

Much of the early quantitative work in the study of gender used a group differences approach and studies exclusively cisgender (those whose gender identity and gender assigned at birth match) men and women. This approach is problematic insofar as it assumes gender differences exist and can exacerbate or create said differences, thus reinforcing the gender binary (see, for example, the controversial example of proficiency in math, Kane & Mertz, 2002; Penner & CadwalladerOlsker, 2012; Unger, 1979; Voyer & Voyer, 2014). This approach is particularly problematic because it is pervasive even where gender is not of primary interest (Westbrook & Saperstein, 2015). Further, this approach often completely ignores transgender people and does not speak at all to non-binary individuals, except to implicitly invalidate and erase them.

A more valuable approach that has been taken is the analysis of gender in terms of patriarchal power structures between men and women (For a discussion of this, see Stewart & McDermott, 2004). This approach often focuses exclusively on cisgender men and women, and is therefore of limited use in understanding transgender individuals, especially those who are male assigned at birth (MAAB, contrasted with female assigned at birth, FAAB; Fleming, Jenkins, and Bugarin, 1980).

A more widely applicable and slightly more inclusive approach to gender is viewing it as a continuum, or two continuums, of masculinity and femininity. In psychology, this approach started with the work of Constantinople (1973), and most commonly used the Bem Sex Role Inventory (BSRI; Sandra L. Bem, 1974). It is also common in lay discourse, as seen in the "Genderbread person" (Figure 1). However, even this approach reinforces the gender binary

(Morawski, 1985) and erases non-binary people. Such an approach implies that there is an objective "masculinity" and "femininity" and that they are each a single coherent entity, despite the existence of countless styles of masculinity and femininity. Genders that exist outside of this entirely cannot be meaningfully studied using existing methods that focus on masculinity and femininity, such as the BSRI. Such methods, by definition, will only ever find masculinity and femininity – thus fundamentally reinforcing the binary.

A Multifaceted Approach

A more inclusive conceptualization of gender is as a multifaceted construct with numerous components, both categorical and continuous, that is related to but distinct from sex. Briefly discussing sex is important because there are many similarities between the conceptual structure of sex, as commonly used in gender psychology and the model of gender proposed in this study. Sex is a multifaceted construct, consisting of chromosomes, primary sex characteristics (internal and external genitalia), secondary sex characteristics (breasts, hair distribution, bone shape, etc.), reproductive capacity, hormone distribution, sex assigned at birth, and current sex category, and is itself socially constructed (See Schellenberg & Kaiser, 2018 or Hyde et al., 2018 for a more extensive overview).

Similarly, gender can be conceptualized as a multifaceted social construct with aspects that are as of yet under explored. Tate et. al. (2014) propose a model with five aspects: birth-assigned gender, current gender identity, gender roles and expectations, gender social presentation, and gender evaluations in their discussion of the study of gender self-categorization and the advantages of integrating the study of cisgender and transgender experiences. These

aspects are a good starting point, but fully understanding gender will require exploring other possible aspects. For example, this study will explore aspects of gender self-schemas.

Further, sex and gender are intimately interrelated, and this is central to a nuanced understanding of the concepts (Schellenberg & Kaiser, 2018). However, the complexity of this model of gender, and the context in which gender research exists requires precision in conceptualization. Developing a comprehensive understanding of gender will require being able to focus on the relevant aspects of sex, gender, and their interplay based on the appropriate context. However, developing this understanding of interrelationships requires first being able to clearly distinguish between them, as failing to do so runs the risk of perpetuating the conflation of sex and gender. Binarist assumptions can influence and be reproduced by research (Schellenberg & Kaiser, 2018; Westbrook & Saperstein, 2015). Because of this and the pervasiveness of the conflation of sex and gender, it is important to highlight the necessity of the conceptual distinction between sex and gender. Transgender, and in particular non-binary gender identities present such a fundamental challenge to our understanding of gender that to properly meet this challenge, the goal of gender research should be to establish a foundational understanding of gender. Given that goal, research would benefit from the conceptual clarity provided by focusing on the study of specific aspects of gender that make clear distinctions between gender and sex, allowing us to begin to understand gender in itself rather than to prematurely add the complexity of interplay between sex and gender. Additionally, most contemporary research outside of gender psychology likely is interested in *either* sex or gender, and could benefit from the conceptual clarity of honing in on the relevant construct. Further, while the interplay of sex and gender is fundamental to transgender experiences of gender, the

transgender community has been, and continues to be harmed by the conflation of sex and gender. Clear conceptual distinctions between sex and gender, and only highlighting interplay when necessary, is important for ensuring the research is correctly understandable by readers less versed in the nuances of gender theory. Further, this clarity and accessibility is important in rectifying the damage our discipline has done in nurturing the culture of transphobia that currently exists. For this reason, the present study will use the terms gender and man/woman rather than sex and male/female throughout the paper regardless of the original language used in the papers discussed.

If psychology as a discipline wishes to understand gender in a fundamental way that is inclusive of all genders, we must adopt new methods. These methods must allow us to ask the question of 'what is gender' without forcing the answer to be limited to one manifestation of it. One way to accomplish this is to focus on studying aspects of gender that could exist and vary within and between all genders. The development of new measures using this approach would allow researchers to begin to meaningfully study gender in a way that accommodates both non-binary individuals and those with expressions of masculinity or femininity outside of what is assessed by previous scales. To highlight the new opportunities this proposed approach presents, and to provide an example of what it might look like in action, this study will demonstrate and validate two measures for studying gender cognition. Further, this study will demonstrate their utility in exploring similarities and differences between different groups of genders. Specifically, I will look at two aspects of people's gender self-schemas: their breadth and salience.

Gender Schematicity

Schemas are cognitive networks of interconnecting concepts that are related to a given category, in this case gender. This topic was chosen because it presents a clear candidate for how individuals or genders may be studied purely in terms of quantitative aspects of schemas, independent of the content of the schemas, and can therefore be relevant across individuals with radically different schema content. I conceptualize one's schematic breadth (SB) as the amount of information contained within the individual's gender schema. At present, to make the scope more manageable, I focused only on an individual's self-schema, which is the information that is considered relevant to an individual's conceptualization of their own gender. This study examines schematic breadth in three components: presentation/performance, role, and identity. Gender presentation is the aspects of one's appearance that they view as gendered, which are chosen to convey an image of a certain gender (or genders) and often to signal one's gender identity. Gender performance is how one "does" gender (Butler, 1990; West & Zimmerman, 1987). It is the behaviours one uses in social interactions that they view as gendered, which serves to further communicate their gender identity or desired gender image. Together, I view these two concepts as representative of the breadth of one's understanding of their gender as something that exists as an image and perception by others. In contrast, gender role consists of socially determined expectations associated with one's gender. This does not necessarily mean that people feel compelled to conform to them, and is not necessarily based on one's sex assigned at birth. Finally, identity breadth is conceptualized as what one views as associated with their own, internal sense of gender, which is independent of presentation or role.

Salience is an individual's tendency to notice stimuli associated with a given concept, or to use those feature(s) to guide the processing of information. For example, an individual high on gender salience would be more likely to judge ambiguous stimuli to be relevant to gender. Essentially, gender salience is the extent to which one uses gender as a lens when viewing the world and the self. I distinguish between salience for self-relevant information, which I term gender identity importance, and the external world, which I term gender schematicity. Gender identity importance is conceptualized as the extent to which one's gender is considered important to their sense of self and something they think about often. Or in cognitive terms, how strongly the self-schema and relevant parts of their gender schema are linked, rather than just the extent of the overlap. Schematicity is the extent to which gender schemas are used to process information about others and the world. It is worth noting explicitly that I view schemas and the processing of schema-relevant information as something which can be conscious. We can be aware of when and how we think of things, and to some extent through conscious effort change how we think about things even absent the conscious effort (for example, learning to use a new set of pronouns for someone and thinking of them as a different gender than before).

The concept of schematicity discussed above builds largely on the work of Sandra Bem (1981; 1982) and her gender schema theory. For Bem, schematicity is the general cognitive readiness to use gender schemas in processing information. Gender schematic individuals are thought to spontaneously sort stimuli into masculine or feminine, and organize their self-concept and behaviour based on this distinction. The critical difference between my approach to schematicity and Bem's is that in her work, schemas are based on the culturally defined roles associated with men and women, whereas I argue that individuals, particularly non-binary people,

can have idiosyncratic gender schemas that do not necessarily bear a strong relation to gender stereotypes. While Bem points out that gender schema theory is a theory of process and not content, her work relies on the BSRI which provides one very limited view of masculinity and femininity that is primarily based on the dichotomies of assertive vs. yielding and instrumental vs. expressive (Bem, 1974; 1981).

Previous Research

There has been relatively little work that examines the salience of identities as an individual differences variable that could inform the proposed study. The Identity subscale of the Collective Self-Esteem Scale (CSES; Luhtanen & Crocker, 1992) provides a measure of how important one's group identity is to them. The scale was built such that it could work for any group, and it is possible to reword the items so that they address gender broadly. However, this requires dropping a number of items as they are irrelevant to gender in general terms. Nevertheless, this is an obvious choice for validating the proposed scale due to its conceptual similarities.

Another measure of identity salience is Callero's (1985) scale for the role-identity salience of blood donors. The scale is somewhat limited for adapting to gender as it is specifically based on identification with a role, but some of the items were adapted for use in this study. The scale itself includes items relevant to identity importance as well as the extent to which one thinks about the identity. As such, it taps into identity importance as conceptualized above in a way beyond the CSES, and was included in the present study for validating the proposed scale.

One of the most flexible and face-valid measures of identity importance is the Twenty Statements Test (TST; Kuhn & McPartland, 1954). In the TST, participants provide twenty spontaneous self-descriptions beginning with "I am…". Variations on this have been used for similar purposes in previous research (e.g., Jackson, 1985). It is necessarily a crude measure of gender identity importance, and often requires coding by researchers which presents a problem for people with idiosyncratic gender schemas. Coding of identity responses is inherently limited to the researcher's conceptualization of gender. A researcher can never fully understand the individual meanings of a gender identity label, particularly one that is highly political or unknown to the researcher, and as such coding risks invalidating the participant and miscategorizing them.

A unique approach to the study of gender salience used in the past is Randel's (2002) measure that focuses on the extent to which one thinks of group composition in terms of gender. This approach maps on neatly to the proposed conceptualization of gender schematicity, however it is inherently binary and modifying the items to be used meaningfully with non-binary people is unwieldy. Regardless, it was included to validate the schematicity subscale.

Additionally, one study found that transsexual women were disproportionately feminine sex-typed compared to cisgender women based on the BSRI (Fleming et al., 1980). This result suggests that the gatekeeping associated with access to transition-related care for TW restricts the range of gender identity/expression, which is expected to be seen in schematic breadth.

Much of this study stems out of my honours thesis, which included an earlier version of the schematic breadth measure, and a single item measure of identity importance (Strazds, 2015).

The study looked at differences between groups of genders on schematic breadth and importance to demonstrate that the measures could be used to meaningfully investigate differences between all types of genders. The data were intended to be compared between cisgender women (CW), cisgender men (CM), transgender women (TW), transgender men (TM), and non-binary individuals (NB). Unfortunately, due to sample size limitations, analysis required collapsing across cisgender (cis)/transgender (trans) status, with three groups including men, women (including both cis and binary transgender (BT) people), and NB. Additionally, analysis was also done grouping across man/woman status, with three groups including cis, BT (including both women and men), and NBs. For schematic breadth, cis people and NB scored higher than BT. For importance, women and NB scored higher than men, and BT scored higher than NB who in turn scored higher than cis people. The present study used the same general approach of testing gender group differences on the measures, and as such the results from the honours thesis form the basis of the present hypotheses related to importance and schematic breadth.

To expand on this previous work and begin to address some of the limitations in the quantitative study of gender, this study will continue to develop these measures of schematic breadth and gender salience. Specifically, schematic breadth will be extended to allow for a more nuanced look at gender schemas through distinctions between different types, and gender identity importance will be developed into a larger gender salience scale. Additionally, this study will begin to validate these measures. For gender salience, this will be done by examining content validity through assessing internal consistency, convergent validity, and to a limited extent discriminant validity. Schematic breadth's content validity will also be examined to a limited extent by exploring discriminant validity.

Methods

Participants and Procedure

The study was conducted through an online survey, hosted by Qualtrics. Participants were recruited through snowball sampling using social media and contacts in the Toronto queer community. The survey was posted in both physical and internet communities, including Toronto and Ontario based communities, and through Reddit.com, specifically the sub-Reddits (a section of the website that is dedicated to a particular purpose or community) r/asktransgender, r/ftm, r/mtf, and r/samplesize (a community for sharing and participating in research studies). This sampling method was required due to the relatively small proportion of the population that are gender minorities. The intent with this sampling procedure was to get a variety of different genders such that meaningful comparisons validating the scale can be made, rather than to claim that it is representative. Inducements were offered in the form of a draw for a $100 cash prize, paid in the form of a gift card.

In total, there were $N = 466$ responses with sufficiently complete data to be used in at least some analyses. By gender group, there were 46 (10.57%) cisgender men, 78 (17.93%) cisgender women, 64 (14.71%) non-binary people, 111 (25.52%) transgender men, 118 (27.13%) transgender women, and 18 (4.14%) who did not wish to be included in any of the prior groups. Most (76.33%) spoke exclusively English and gave easily parsed responses. Most participants came from the USA (55.48%), Canada (26.11%), and the UK (7.93%).

Measures

The following measures were used in the study. Unless stated otherwise, the order used here were the order of presentation.

Analytic Category. Participants were asked which category they wished to be placed in for purposes of statistical analysis. They were presented with the prompt "Unfortunately, due to limitations on the sample sizes that can reasonably be collected, participants will need to be grouped into different types of genders for the purposes of statistical analysis. Which group do you feel is most appropriate for your gender? Additionally, feel free to leave feedback about this approach in the text box below." The options were "Cisgender Woman", "Cisgender Man", "Transgender Woman", "Transgender Man", "Non-Binary", and "I do not wish to be placed in any of these groups." If participants chose the latter, they were presented with another question asking why they chose to be excluded and if they have any suggestions for how they could be accommodated. A text box for feedback was also displayed at the bottom of the page. This question was included after participants had completed other substantive questions, after a more inclusive and extensive gender question and before other demographic questions.

Schematic Breadth (SB). Participants were shown the following prompts, and filled in their answers in an open-ended format in a text box. Schematic breadth is the total number of terms provided in response to all prompts. The prompts are viewed as measures of gender expression breadth, gender role breadth, and gender identity breadth, respectively.

"Please list all the words or phrases that are related to your gender presentation/expression or your gender performance. Gender presentation and expression means the aspects of your

appearance that are experienced as gendered by you. Gender performance means the behaviours or ways of interacting with people that are experienced as gendered by you. For clarity, please separate each word or phrase with a semicolon (;).”

“Please list all the words or phrases that are related to society’s expectations for your gender. Please note that this question is not asking about expectations for your sex assigned at birth, or expectations that you necessarily feel you need to meet, but rather any expectations you are aware of that are relevant to your gender. For clarity, please separate each word or phrase with a semicolon (;).”

“Please list all the words or phrases that are related your own internal sense of gender. For clarity, please separate each word or phrase with a semicolon (;).”

Verbal Fluency (VF). Participants were given one minute to list as many words as possible that start with the letter R, using a semicolon break between words as in the schematic breadth prompts. Their score is the number of unique words. This will allow controlling for one’s ability to spontaneously generate words not necessarily related to gender in analyses that include schematic breadth.

Gender Salience. Participants rated the following items on a sliding scale from 0 to 6, with floating labels from “Very Strongly Disagree”, “Strongly Disagree”, “Disagree Somewhat”, “Neither Agree nor Disagree”, “Agree Somewhat”, “Strongly Agree”, and “Very Strongly Agree”. This scale is broken down into two subscales: identity importance and schematicity.

Table 1

Gender Salience Scale Items

Gender Identity Importance (GII)
"My gender identity is important to me."
"I frequently think about my gender identity."
"My gender is an important part of who I am."
"My gender is an important part of my sense of self."
"My gender is an important part of my sense of self."
"My gender is an important part of my sense of self."
"I spend a lot of time thinking about what my gender means to me."
Gender Schematicity (GSC)
"I often wonder what other people think about their gender"
"I pay attention to what pronouns people use."
"I notice or think about how much space that I take up relative to other genders in group conversations."
"Gender is an interesting thing to think about."

Validity Measures

Twenty Statements Test (TST). Participants completed the TST (Kuhn & McPartland, 1954), in which they list up to twenty words that complete the prompt "I am…". To score responses for gender salience, rather than traditional scoring, another question was included in which participants rate each item in response to the prompt "Earlier you listed a number of terms that describe yourself. Please rate them on how relevant they are to your gender. If you did not fill in all twenty items earlier in the study, there will be some blank sections. Please disregard these." They were rated on a sliding scale from 0 to 6, with floating labels of "Not at all relevant", "Moderately relevant", and "Very relevant". The total score is the sum of the relevance score multiplied by the inverse of its rank (e.g., item number 20 is 1, item number 1 is 20). Weighting items listed first more heavily than later items allows the score to reflect both how strongly one considers aspects of their identity to be relevant to their gender, but also how salient those aspects are, in terms of being listed first. This was the first measure participants completed.

Role-Identity Salience. Participants completed a modified version of Callero's (1985) measure of salience. The following items were rated on a 9 point Likert-type scale from strongly disagree to strongly agree. One item was omitted on the basis of not being applicable to gender as conceptualized in this study.

Table 2

Modified Callero's Role-Identity Salience Scale Items

"Gender is something I rarely even think about." (Reverse scored)
"I really don't have any clear feelings about gender." (Reverse scored)

"For me, my gender is more than just an identity."
"Gender is an important part of who I am."

Collective Self-Esteem Scale. Participants completed a modified version of the identity subscale from Luhtanen & Crocker's (1992) Collective Self-Esteem Scale. The following items were rated on a 7 point Likert-type scale from strongly disagree to strongly agree.

Table 3

Modified Collective Self-Esteem Scale Items

"Overall, my gender has very little to do with how I feel about myself." (Reverse scored)
"My gender is an important reflection of who I am."
"My gender is unimportant to my sense of what kind of person I am." (Reverse scored)
"In general, my gender is an important part of my self-image."

Identity Salience. Participants completed Randel's Identity Salience scale (2002). The following items were rated on a 5 point Likert-type scale from strongly disagree to strongly agree.

Table 4

Modified Randel's Identity Salience Scale Items

"When people ask me about who is in the group, I initially think about describing the group in terms of gender composition (e.g., two women and three men)."

"It is not intentional, but when I think of my fellow group members, what comes to mind initially is the names of the women then the names of the men (or the men's names and then the women's names)."

"Even though I don't mean to, I think of their gender as the most prominent characteristic of my fellow group members."

Transphobia. Participants completed Nagoshi et al.'s (2008) transphobia scale. The scale contains 9 items and measures anti-transgender prejudice on a scale of 1 to 7 from "Completely Disagree" to "Completely Agree". Example items include "I don't like it if someone is flirting with me, and I can't tell if they are a man or a woman.", "I think that there is something wrong with someone who says that they are neither a man nor a woman.", and "I would be upset, if someone I'd known for a long time revealed to me that they used to be another gender."

Social Desirability. Participants completed Bobbio and Manganelli's (Bobbio & Manganelli, 2011) short form of the Balanced Inventory of Desirable Responding (BIDR; Paulhus, 1991), which measures socially desirable responding through two subscales including Self-Deceptive Enhancement (SDE) and Impression Management (IM). The measures are on a 6 point scale, where higher scores represent a greater tendency towards socially desirable responding. Example items include "I always know why I like things", "I am a completely rational person", "I always obey laws, even if I'm unlikely to get caught", and "I have some pretty awful habits (reverse scored)".

Demographics. Finally, participants completed a number of demographic questions, including age, sex, sex assigned at birth[2], gender identity, sexual orientation, relationship status, living arrangements, religion, country of birth, what country they live in, ethnic identity, and how they found out about the study.

Hypotheses

Schematic breadth was expected to be similar between CW and CM, as well as TM and NB, and both groups will differ from each other and TW, based on my previous research (Strazds, 2015). Further, it was expected that the proportion of different parts of schematic breadth will differ, with CM and CW using more role terms, while TW, TM, and NB use more identity and expression terms. For identity importance by gender, it is difficult to hypothesize an exact pattern of gender differences, however CM are expected to score lower than all other genders. Finally, it was expected that importance would be positively related to the TST (Kuhn & McPartland, 1954) Callero's (1985) salience measure, and Luhtanen & Crocker's (1992) identity subscale. The schematicity subscale was expected to be positively related to Randel's (2002) salience scale.

[2] Note that while sex and sex assigned at birth were collected in this study, this should not be taken as an example of how to collect gender information for psychological research. Unless needed for a specific research question about sex or gender histories, which would regardless prompt more nuanced questions than the ones used here, these two questions are best excluded.

Results

Scale Validation

Reliability

The gender identity importance (GII) subscale had a high reliability, α = .874, SE = .009. However, one item that was initially collected ("I would be upset if someone thought I was a gender other than the one I identify with") was noted as having a comparatively low *r* with the rest of the subscale, *r* = .451, and was not included in the final scale based on theoretical considerations (see discussion). The gender schematicity (GSC) subscale had a decent reliability, α = .769, SE = .017. One item ("It is important to know people's gender") was excluded as the scale was considerably less reliable with it included, α = .695, SE = .022, and had a very low *r* with the rest of the subscale, *r* = .117. Finally, the overall scale had a high reliability, α = .864, SE = .00

Table 5

Correlation Matrix of Main Measures

	GII	GSC	SBR	SBP	SBI	VF	TST	Randel	Callero	CSES	TS	SDE	IM
GII	1	.516	.081	.217	.236	.003	.457	.069	.745	.648	-.236	-.018	-.097
GSC	.516	1	.133	.214	.207	.012	.274	.025	.439	.305	-.398	-.192	-.049
SBR	.081	.133	1	.438	.445	.182	.059	-.008	.108	.057	-.068	-.144	-.006
SBP	.217	.214	.438	1	.548	.191	.123	-.097	.196	.167	-.167	-.11	-.059
SBI	.236	.207	.445	.548	1	.191	.203	-.039	.254	.195	-.08	-.047	.004
VF	.003	.012	.182	.191	.191	1	-.144	-.135	.002	.083	-.035	-.142	.016
TST	.457	.274	.059	.123	.203	-.144	1	.081	.431	.363	-.176	.015	-.004
Randel	.069	.025	-.008	-.097	-.039	-.135	.081	1	.039	.164	.292	-.013	-.03
Callero	.745	.439	.108	.196	.254	.002	.431	.039	1	.648	-.283	-.059	-.056
CSES	.648	.305	.057	.167	.195	.083	.363	.164	.648	1	-.107	-.009	-.071
TS	-.236	-.398	-.068	-.167	-.08	-.035	-.176	.292	-.283	-.107	1	.138	.013
SDE	-.018	-.192	-.144	-.11	-.047	-.142	.015	-.013	-.059	-.009	.138	1	.103
IM	-.097	-.049	-.006	-.059	.004	.016	-.004	-.03	-.056	-.071	.013	.103	1

Note. Used Spearman's Rank-order correlation. Acronyms: Gender Identity Importance (GII), Gender Schematicity (GSC), Schematic Breadth Role (SBR), Presentation (SBP), Identity (SBI), Verbal Fluency (VF), Collective Self-Esteem Scale (CSES), Transphobia scale (TS), Self-Deceptive Enhancement (SDE), Impression Management (IM)

Validity

Correlations between the gender salience subscales and other salience measures were conducted to confirm the content validity of the scale (See Table 5 for correlation matrix of all validity variables). Many of the variables were non-normally distributed, so the Spearman rank-order correlation was used. As predicted, gender identity importance was significantly correlated with TST gender relevance ratings, r_s= .457, p < .001, demonstrating convergent validity. Similarly, gender identity importance was strongly correlated with the CSES, r_s = .648, p = .002. Finally, gender identity importance was strongly correlated with Callero's salience measure, r_s= .745, p < .001. However, gender schematicity was not significantly correlated with Randel's salience measure, r_s = .025, p = .606.

In order to investigate the discriminant validity of the measures created in this study, regressions were conducted with transphobia and the BIDR as predictors. The schematic breadth regressions were conducted using a zero-truncated negative binomial distribution, given the count nature of the data, and some indication of over dispersion. For gender identity importance and schematicity, there was no indication of serious non-normality, heteroscedasticity, or problematic outliers. Only SDE significantly negatively predicted the role component of schematic breadth, however the effect is small, with each point predicting a difference of only roughly one tenth of a term (See Table 6). There were no significant predictors for the identity component of schematic breadth (See Table 7). Only transphobia negatively predicted the presentation component of schematic breadth, however it was a similarly weak effect, with each point predicting a difference of only one tenth of a term (See Table 8). Identity importance was significantly negatively

predicted by transphobia and IM, with a modest effect size, as the whole model accounted for less than 9% of the variance in terms used (See Table 9). Finally, transphobia and SDE significantly negatively predicted gender schematicity, and transphobia in particular large effect, with each point of transphobia predicting a difference of over half a point in gender schematicity (See Table 10).

Table 6

Schematic Breadth (Role) Zero-truncated Negative Binomial Regression Table

	Estimate	Std. Error	z	*p*
Intercept 1	2.16	0.2	10.75	< .001
Intercept 2	1.13	0.14	7.86	< .001
Transphobia	0	0.04	-0.05	.962
SDE	-0.12	0.05	-2.59	.009
IM	-0.02	0.04	-0.38	.707

Note. Log-likelihood: -988.45, df: 765

Table 7

Schematic Breadth (Identity) Zero-truncated Negative Binomial Regression Table

	Estimate	Std. Error	z	*p*
Intercept 1	1.62	0.25	6.38	< .001
Intercept 2	0.72	0.18	3.94	< .001
Transphobia	-0.04	0.05	-0.73	.466
SDE	-0.1	0.06	-1.68	.093
IM	0	0.05	0.05	.956

Note. Log-likelihood: -844.64, df: 759

Table 8

Schematic Breadth (Presentation) Zero-truncated Negative Binomial Regression Table

	Estimate	Std. Error	z	*p*
Intercept 1	2.13	0.24	9.06	< .001
Intercept 2	0.76	0.15	5.07	< .001
Transphobia	-0.14	0.05	-2.82	.005
SDE	-0.09	0.05	-1.77	.078
IM	-0.01	0.05	-0.15	.882

Note. Log-likelihood: -968.25, df: 773

Table 9

GII Regression Table

Predictor	β	95% CI	$t(421)$	p
Intercept	5.20	[4.51, 5.89]	14.79	< .001
Transphobia	-0.40	[−0.53, −0.26]	-5.87	< .001
BIDRSDE	0.06	[−0.10, 0.21]	0.70	.482
BIDRIM	-0.15	[−0.30, 0.00]	-2.01	.045

Note. $(3, 421) = 13.25$, $p < .001$, $r^2 = .086$.

Table 10

GSC Regression Table

Predictor	β	95% CI	$t(420)$	p
Intercept	5.83	[5.22, 6.44]	18.75	< .001
Transphobia	-0.60	[−0.72, −0.49]	-10.17	< .001
BIDRSDE	-0.21	[−0.34, −0.07]	-2.96	.003
BIDRIM	0.02	[−0.11, 0.15]	0.27	.789

Note. $(3, 420) = 40.97$, $p < .001$, $r^2 = .226$.

GII & GSC by Gender

To demonstrate the potential utility of this gender salience scale in studying gender, differences in gender salience subscales were explored between gender groups. For both gender identity importance (GII, Figure 2 & 3) and gender schematicity (GSC, Figure 4 & 5) distributions deviated considerably from normality, and from each other both in terms of

distribution shape and variance. Therefore, 20% trimmed means were used for conducting Analysis of Variance (ANOVA)s between gender groups (Wilcox, 2012).

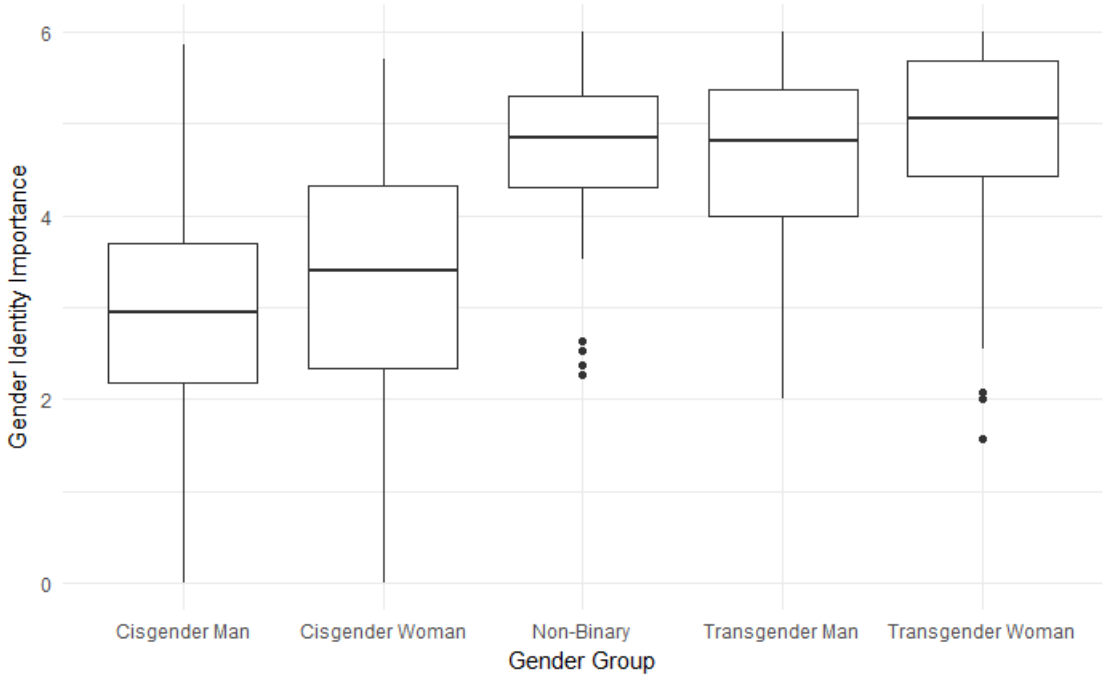

Figure 2 Boxplot of GII by Gender

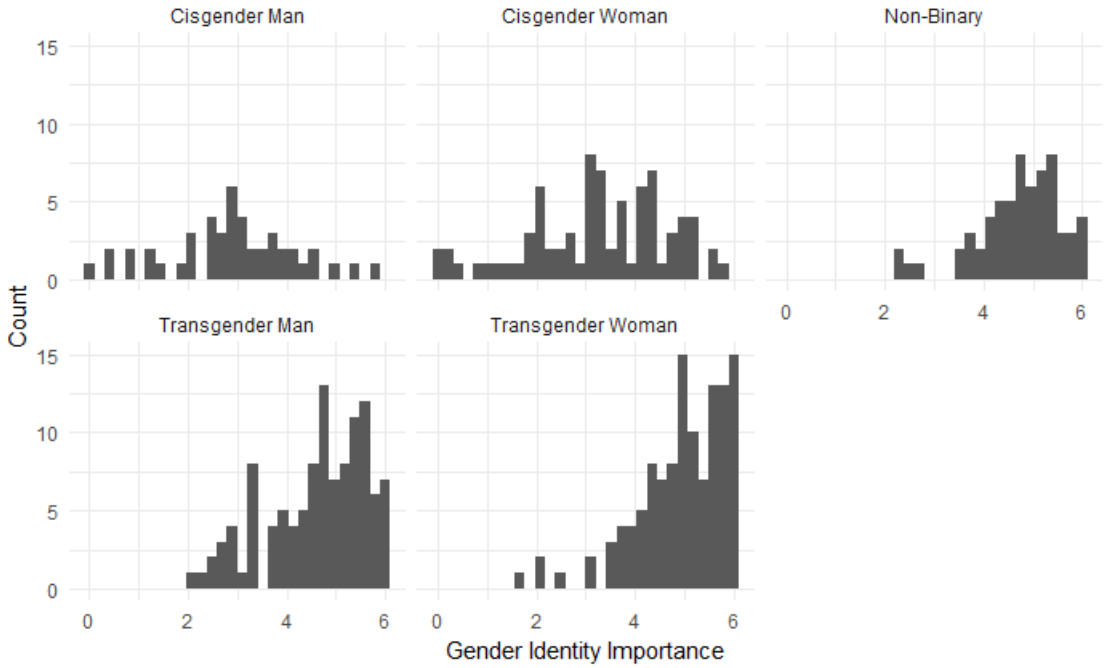

Figure 3 Histogram of GII by Gender

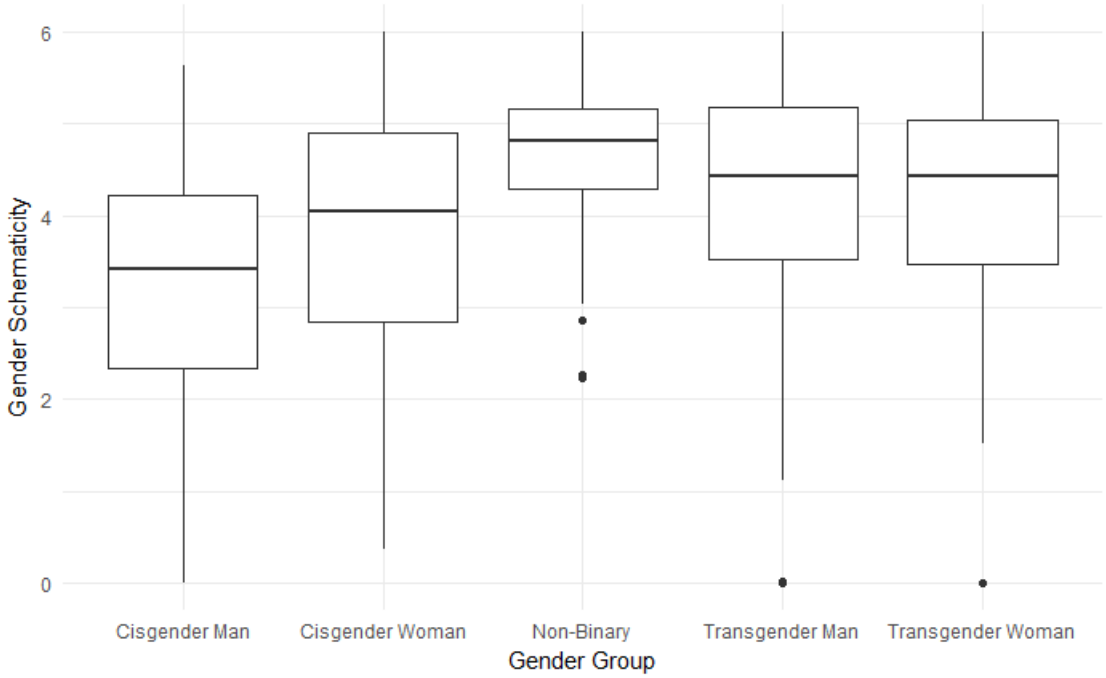

Figure 4 Boxplot of GSC by Gender

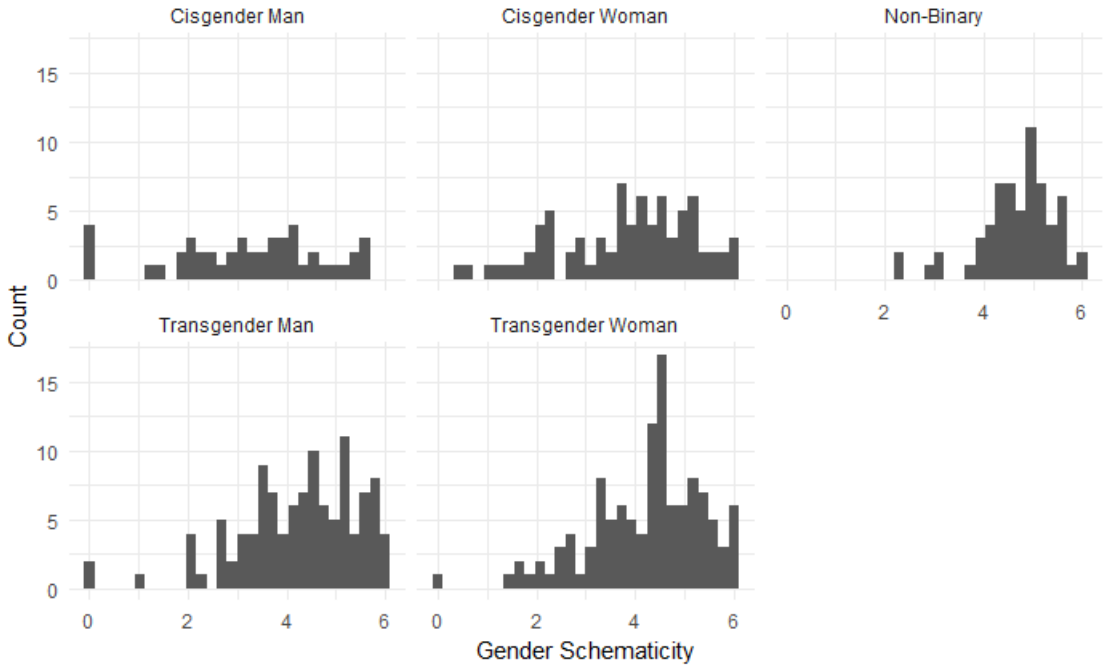

Figure 5 Histogram of GSC by Gender

Identity importance significantly differed between gender groups $F(4, 102.39) = 42.49$, $p < .001$. All pairwise comparisons were tested. (See Table 11) The results show a clear trend of cisgender people scoring significantly lower than transgender people. These results support the hypothesis that CM would score lower than others for trans people, but are ambiguous whether they differ from CW. Further, it is ambiguous whether TW scored significantly higher than TM and NB.

Table 11

Pairwise Comparisons for ANOVA of GII by Gender

Comparison	$\hat{\Psi}$	95% CI	*p*
Cisgender Man vs. Cisgender Woman	-0.42	[-1.13, 0.3]	.097
Cisgender Man vs. Non-Binary	-1.86	[-2.46, -1.26]	< .001
Cisgender Man vs. Transgender Man	-1.82	[-2.43, -1.22]	< .001
Cisgender Man vs. Transgender Woman	-2.12	[-2.7, -1.54]	< .001
Cisgender Woman vs. Non-Binary	-1.44	[-2.02, -0.86]	< .001
Cisgender Woman vs. Transgender Man	-1.41	[-1.99, -0.82]	< .001
Cisgender Woman vs. Transgender Woman	-1.71	[-2.26, -1.15]	< .001
Non-Binary vs. Transgender Man	0.04	[-0.39, 0.46]	.815
Non-Binary vs. Transgender Woman	-0.26	[-0.65, 0.12]	.056
Transgender Man vs. Transgender Woman	-0.3	[-0.69, 0.1]	.035

Note. Based on 20% trimmed means. *p* values are not corrected for multiple comparisons, however confidence intervals are.

Likewise, gender schematicity significantly differed between gender groups $F(4, 102.88) = 10.99$, $p < .001$. All pairwise comparisons were tested (See Table 12). The general pattern was NB scored highest, followed by TM and TW, then CW, and CM scored lowest. Whether differences are statistically significant is somewhat ambiguous between groups within the aforementioned hierarchy, and the precise rankings may not be possible to replicate.

Table 12

Pairwise Comparisons for ANOVA of GSC by Gender

Comparison	$\hat{\Psi}$	95% CI	*p*
Cisgender Man vs. Cisgender Woman	-0.59	[-1.47, 0.29]	.056
Cisgender Man vs. Non-Binary	-1.42	[-2.17, -0.67]	< .001
Cisgender Man vs. Transgender Man	-1.03	[-1.81, -0.25]	< .001
Cisgender Man vs. Transgender Woman	-0.99	[-1.76, -0.22]	< .001
Cisgender Woman vs. Non-Binary	-0.83	[-1.43, -0.23]	< .001
Cisgender Woman vs. Transgender Man	-0.44	[-1.09, 0.21]	.056
Cisgender Woman vs. Transgender Woman	-0.39	[-1.03, 0.24]	.078
Non-Binary vs. Transgender Man	0.39	[-0.03, 0.81]	.01
Non-Binary vs. Transgender Woman	0.43	[0.04, 0.83]	.002
Transgender Man vs. Transgender Woman	0.05	[-0.42, 0.51]	.785

Note. Based on 20% trimmed means. *p* values are not corrected for multiple comparisons, however confidence intervals are.

SB by Gender & Type

To demonstrate the nuance this approach can achieve in studying gender, differences in schematic breadth scores across gender groups (CM, CW, NB, TM, TW) and type of prompt (role, presentation, identity) were analyzed. A mixed multilevel modelling approach to accommodate dependence across type of prompt. The number of schematic breadth terms used by the participant was chosen as the outcome variable to allow for examining differences in number of terms used across types. As such, term type was included as predictors within the

model, and the effect of type within participant was included as a random slope. To explore gender differences in term usage across types, gender group was included as predictors. Additionally, to explore proportional differences in term usage across type and gender, interaction terms were included in the model. Finally, verbal fluency was included as a covariate. Due to the count nature of the data and no concerning evidence of over dispersion, the model used a zero-truncated Poisson distribution.

A hierarchical regression was conducted using each variable and gender by type interaction term successively (See Table 13). Only verbal fluency, prompt type, and gender by type interaction were significant. Contrary to my hypotheses, there was no significant main effect of gender.

Table 13

Hierarchical Regression for SB Gender by Type Model

	df	AIC	BIC	Log Likelihood	Deviance	χ^2	χ^2df	p
Base	3	5304.67	5319.74	-2649.34	5298.67			
VF	4	5290.66	5310.75	-2641.33	5282.66	16.01	1	< .001
Type	6	5187.09	5217.22	-2587.54	5175.09	107.57	2	< .001
Gender	10	5187.83	5238.05	-2583.91	5167.83	7.26	4	.123
Gender x Type	18	5130.29	5220.68	-2547.14	5094.29	73.54	8	< .001

To ensure that no gender or perspective was unduly privileged by being treated as a reference group, and to better understand the trends given multiple interactions, the analysis was redone rotating through each possible reference group and prompt type. To effectively and efficiently

report such numerous statistics, I will highlight the overall patterns within the data and deviations from them, providing representative values where it is possible to do so concisely or where the results are comparatively statistically ambiguous. Interested readers should consult the full tables for further detail (see Tables 14 – 28 in Appendix A). For an overview, a summary table of gender comparisons for each prompt type is also provided (See Tables 29 - 31), as well as a boxplot of schematic breadth scores by gender at type (See Figure 6). It should be noted that given the nature of this analysis necessitating so many comparisons to get a full view of the data, it was deemed overly conservative to apply corrections to *p* values. Given the number of comparisons it is almost certain corrections would suppress any of the more subtle differences in the patterns. Further, given that the design of this analysis means replicating any aspect of it would likely necessitate replicating the whole analysis, the risks of false-positives do not seem as drastic as the risk of over correction in exploratory research. As such, the following interpretation is based not on strict cutoffs of significance based on *p* values, but rather by considering the *p* values in the context of the relevant other trends and comparisons. I endeavor to highlight areas where the results are less certain. Regardless, these results should be interpreted with caution and any particular detail should be considered tentative.

Table 29

Schematic Breadth Poisson Multilevel Model Summary Table for Presentation

Comparison	β	Std. Error	z	*p*
CW vs CM	0.39	0.16	2.44	.015
NB vs CM	0.7	0.16	4.33	< .001
TM vs CM	0.28	0.15	1.82	.069
TW vs CM	0.36	0.15	2.35	.019
NB vs CW	0.31	0.13	2.49	.013
TM vs CW	-0.11	0.11	-0.98	.325
TW vs CW	-0.03	0.11	-0.26	.795
TM vs NB	-0.42	0.12	-3.63	< .001
TW vs NB	-0.34	0.12	-2.9	.004
TW vs TM	0.08	0.1	0.78	.435

Table 30

Schematic Breadth Poisson Multilevel Model Summary Table for Role

Comparison	β	Std. Error	z	*p*
CW vs CM	0.24	0.14	1.69	.091
NB vs CM	-0.17	0.16	-1.1	.27
TM vs CM	0.15	0.14	1.12	.264
TW vs CM	0.12	0.14	0.83	.407
NB vs CW	-0.42	0.13	-3.18	.001
TM vs CW	-0.09	0.11	-0.84	.402
TW vs CW	-0.13	0.11	-1.18	.24
TM vs NB	0.33	0.12	2.65	.008
TW vs NB	0.29	0.12	2.3	.021
TW vs TM	-0.04	0.1	-0.38	.701

Table 31

Schematic Breadth Poisson Multilevel Model Summary Table for Identity

Comparison	β	Std. Error	z	*p*
CW vs CM	0.02	0.17	0.12	.903
NB vs CM	0.41	0.17	2.41	.016
TM vs CM	0.02	0.16	0.14	.887
TW vs CM	0.32	0.16	1.98	.048
NB vs CW	0.39	0.14	2.81	.005
TM vs CW	0	0.13	0.02	.988
TW vs CW	0.29	0.12	2.36	.018
TM vs NB	-0.39	0.13	-3.07	.002
TW vs NB	-0.1	0.12	-0.77	.44
TW vs TM	0.29	0.11	2.62	.009

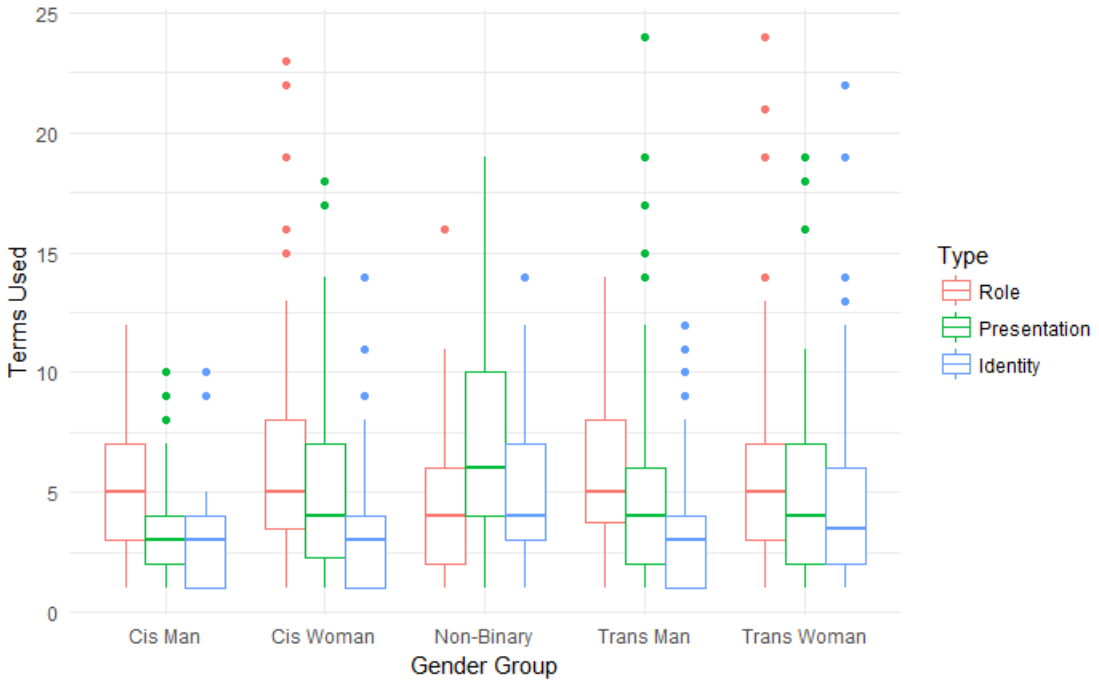

Figure 6 Boxplot of SB by Gender and Type

Note. 2 outliers over 25 were removed to improve scaling.

Overall, CW and TM used significantly more role terms than presentation, and more presentation than identity. TW roughly follow this pattern, however their predicted scores vary less by type and as such the significance of the differences is less certain (Compare Tables 15 & 17 to Table 18). CM differed from this slightly insofar as presentation and identity did not significantly differ, (β = -0.11, p = .451). NB also differed from this pattern. While predicted identity counts are lower than presentation (β = -0.40, p < .001), so are role (β = -0.49, p < .001). Predicted role and identity counts did not differ (β = -0.10, p = .335). This can be seen as following the pattern for presentation vs. identity, but with deflated role terms.

The trend by gender within role can be best understood by focusing on NB. In absolute terms, NB used fewer role terms than CW (β = 0.42, p = .001), TM (β = 0.33, p = .008), and TW (β = 0.29, p = .021), but not compared to CM (β = 0.17, p = .27), however used proportionally fewer role terms than presentation or identity terms compare to all other genders. Otherwise, there were no significant differences between other genders.

The trend within presentation by gender is such that CW, TW, and TM do not differ significantly, but NB used more terms than all others and CM used fewer than others. However the trend of CM using fewer terms is less certain with CW (β = 0.39, p = .015) and TW (β = 0.36, p = .019), and is somewhat ambiguous with TM (β = 0.28, p = .069). Further, compared to CW, CM used proportionally fewer presentation than identity terms (β = 0.37, p = .036).

Within identity, TW and NB used more terms than others. Within those two groups there were no significant differences, however TW use proportionally more identity terms than presentation terms compared to NB. Further, compared to CW and TM, TW use proportionally more identity terms than role and presentation terms.

It is difficult to say precisely whether my hypothesis of trans people using proportionally more identity and presentation terms, and cis people use more role terms was supported. While some comparisons between groups supported it, such as NB using fewer role terms, others were not, such as CW often following a pattern similar to that of TM and TW. Ultimately, these results paint a much more nuanced picture than originally predicted.

Discussion

The measures developed and examined in this study show promise for expanding the scope of gender research in psychology, however further work is required to refine them. Within the gender salience scale, identity importance demonstrated good reliability. Overall the subscale had a good Cronbach's α. The one item that did not fit, "I would be upset if someone thought I was a gender other than the one I identify with" does not fit as well conceptually as the rest of the items. Arguably the item could have different weight and meaning to different genders, who have had different experiences with how their gender is perceived or treated. As such, it is in some sense an "are you trans" question, which could present a confound for research between genders. Further, the pattern of correlations with related measures is encouraging and supports the content validity of the measure. All of the predicted measures had a modest to strong correlation with gender identity importance, and the differences in the strength of the relationship between various measures is informative. The middling-strength correlation with TST ratings could be due to taking a different approach to measurement than the traditional Likert-type scales used in Callero's measure and gender identity importance, both of which had similar correlations to TST ratings, which suggests the measures are circling the same concept but with different methods. Finally, the significant but weak relationship to transphobia supports the discriminant validity of gender identity importance. While some relationship is to be expected insofar as trans people score higher than cis people on identity importance, the weak relationship demonstrates that they are nevertheless distinct concepts. Overall, these results show that this measure provides a good conceptual foundation for further development of a more refined scale.

The results for gender schematicity are less clear. While there was an adequate Chronbach's α, the item that did not fit ("It is important to know people's gender") does not have as neat an explanation as in identity importance. It could be that the item is comparatively politically charged, as gender minorities may be more aware of the relevance of gender, or people may actively avoid explicitly considering gender for fear of appearing sexist. On a similar note, the items seem to be split conceptually into two groups: pure schematicity ("I often wonder what other people think about gender" & "Gender is an interesting to think about"), and awareness of gender issues ("I pay attention to what pronouns people use" & "I notice or think about how much space I take up relative to other genders in group conversations"). In the case of the latter, being cognizant of the importance of not assuming gender identities, and the gendered dynamics in conversations seem to address an understanding of gender politics, for which gender schematicity is a necessary but not sufficient condition. As such it is a distinct, albeit related concept and should not be conflated with gender salience. Further, it is something that would be expected to be related to gender, for example trans people's experiences with misgendering highlights the need to ask for pronouns, and women are more likely to be aware of men talking more in conversations. Therefore, it presents a confound for being able to study the effect of gender schematicity between gender groups, and should be treated as a distinct concept. It is unclear why there was no relationship between schematicity (or any of the salience measures) and Randel's scale. Perhaps the group lens is not appropriate for this conceptualization of schematicity, or it taps in to something else entirely. Additionally, while the modest relationship with transphobia supports discriminant validity, but that it was more strongly related to gender schematicity than identity importance highlights the problem of a gender politics confound within

schematicity. Regardless, this construct clearly requires a more thoughtful approach in subsequent scale development, and the results for schematicity presented here should be interpreted with caution.

That among schematic breadth types transphobia was only significantly related to the presentation component is somewhat surprising, however it does largely fit conceptually with transphobia as disgust towards deviation from gender as an immutable binary. In this context, presentation is the area in which visible transgression from strict social norms rooted in the binary can happen. It is unsurprising that the role component was unrelated, as it is usually composed of socially-prescribed norms within the binary. Similarly, it is somewhat surprising that identity was unrelated to transphobia, as one might expect broader identity schemas to stem from departing from the limitations of the binary. This is particularly surprising given that identity was highest among TW & NB, who are often the more visibly transgressive genders. However, it is possible a more targeted or qualitative analysis of this data could provide more information.

The pattern of gender salience subscales across gender groups is intuitive, however some of the details are ambiguous. For identity importance, it makes sense that trans people, who often have to fight for their identities to be taken seriously, would score higher on identity importance. Further, there was an ambiguous potential trend of women scoring higher than others - specifically CW potentially scoring higher than CM and TW potentially scoring higher than other trans people. For TW, this could be in part due to needing to fight transmisogyny and institutional gatekeeping in access to care, similar to the findings of Fleming (1980). Regardless, it seems that

the most important difference is between cis and trans people, and there is possibly another, smaller effect among women, particularly trans women. Further, this pattern mirrors what I found in my honour's thesis, particularly considering the binary trans grouping was predominantly TW (Strazds, 2015).

While the precise details of the pattern of schematicity across gender are statistically uncertain, the overall pattern has an intuitive theoretical explanation. People who are more exposed to gender-based discrimination would reasonably be more likely to see gendered dynamics at play in their lives, and therefore be more accustomed to using a gendered lens when viewing the world. However, related research in the context of race has been somewhat mixed (Hurtado, Alvarado, & Guillermo-Wann, 2015; Thompson, 1999). Additionally, the confound with gender politics also helps explain this pattern; for many trans people gender can't necessarily be readily perceived, especially NB whose gender is likely invisible and non-normative. Further, it makes sense that CM would score the lowest, as those who are least negatively impacted by gender-based discrimination have less incentive to be aware of it and have fewer opportunities to learn about it. However, this measure needs considerable further development as the results are statistically ambiguous, so this interpretation should be explored further with a more refined measure.

While it wasn't tested directly beyond the hierarchical regression, the lack of any significant differences between genders for overall schematic breadth is worth discussing, as it does not replicate my honours thesis findings or match my hypotheses. This is likely due to the change to measuring schematic breadth as three types, as it is now more nuanced. For example,

NB scored considerably higher on presentation and identity, and role is arguably less conceptually relevant for them, so comparing overall schematic breadth for them may not be as appropriate as in the single-prompt version of the measure. This version provides a much more nuanced approach and more informative findings when considering differences in type and proportions between genders.

Examining schematic breadth as three distinct types, and how the proportions vary across genders provides an interesting demonstration of how this approach to studying gender can lead to a deeper and more inclusive understanding of gender than binary-centric approaches can. While the full analysis presented a very detailed picture, I will focus the discussion on the larger pattern within the results, as many of the findings are comparatively tangential and statistically ambiguous. More targeted follow-up research should be done to explore them more fully.

Overall, the pattern of type proportions is that schemas are composed primarily of role terms, followed by presentation and then identity terms. This pattern, and the major ways different genders deviate from it is both theoretically intuitive, and insightful. The ordering of role - presentation - identity also follows a pattern from socially-prescribed and external to the self, to possibly self-defined and internal to the self. This is in its own right a potentially interesting avenue for future research, but it also informs the differences between genders. Notably, NB people clearly deviated from this pattern. The pattern of using fewer role terms and comparatively more presentation and identity terms fits the interpretation that for many NB people, their genders are based less on social norms, since most non-binary genders exist outside of, or in resistance to those societal gender norms. Conversely, not having a script to follow (at

least in contemporary, white, western societies) means that they are required to explore and define their gender for themselves, therefore resulting in broader or more nuanced understandings of their gender. It seems likely that for them, the concept of gender role simply doesn't mean the same thing as it does when one talking about binary genders which is reflected in terse and idiosyncratic responses often more akin to the identity responses.

This interpretation neatly highlights the importance of inclusive conceptual frameworks when studying gender. By searching for and studying aspects of gender that are applicable to *all* genders, the research is not only more inclusive but can challenge our assumptions about gender. Much of the research on gender in psychology has focused largely on gender roles, but this research suggests a much more nuanced approach is required to understand gender roles in the context of non-binary people, and a more multifaceted approach to the study of gender is warranted.

Limitations & Future Research

Demographics present an obvious limitation to this study. While the intent in sampling was to have a diverse rather than representative sample, it is still important to keep the limitations of the sampling methods in mind when interpreting the results. The most significant example of this is the lack of participants who identified with non-western gender identities. Just as non-binary genders present a challenge to the conceptual framework of gender within psychology, so do non-western genders. This study aims to present an approach to the study of gender proposed in this study that can be inclusive of all genders. Further research is required to consider how

these measures apply to non-western gender identities, and how these identities might inform or challenge our understandings of gender.

Results based on the gender salience subscales developed in this study should be interpreted with caution, and should not be used for further research in their current form. While the identity importance subscale shows promise given further development, the schematicity subscale must be reconsidered to ensure it has adequately clear conceptualization and face validity. While these results will hopefully inspire new approaches to the study of gender, the measures themselves require more extensive pre-planning and validation before they can be recommended for use in further research. A much more extensive approach to scale development and validation is required for gender salience, as this study presented a very limited look at internal consistency, and discriminant validity, and did not begin to assess external validity.

Schematic breadth could be further validated by beginning to assess external validity by investigating potential correlates with meaningful outcomes, such as measures of mental well-being in trans people to assess criterion validity. Additionally, investigating the processes that lead to broader or more nuanced gender schemas could help validate the theoretical foundation of the measure, and lead to interesting avenues for further research.

References

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 42*(2), 155–162.

Bem, S. L. (1981). Gender schema theory: A cognitive account of sex typing. *Psychological Review, 88*(4), 354.

Bem, S. L. (1982). Gender schema theory and self-schema theory compared: A comment on Markus, Crane, Bernstein, and Siladi's' Self-schemas and gender.

Bobbio, A., & Manganelli, A. M. (2011). Measuring social desirability responding. A short version of Paulhus' BIDR 6. *Testing, Psychometrics Methodology in Applied Psychology, 18*(2), 117–135.

Bornstein, K. (1992). *Gender outlaw: On men, women, and the rest of us.* New York: Routledge.

Butler, J. (1990). *Gender Trouble: Feminism and the Subversion of Identity*. New York: Routledge.

Butler, J. (1993). *Bodies that matter: On the discursive limits of sex*. New York: Routledge.

Callero, P. L. (1985). Role-Identity Salience. *Social Psychology Quarterly, 48*(3), 203. doi:10.2307/3033681

Constantinople, A. (1973). Masculinity-femininity: An exception to a famous dictum? *Psychological Bulletin, 80,* 389–407.

Corwin, A. I. (2009). Language and gender variance: Constructing gender beyond the male/female binary. *Electronic Journal of Human Sexuality, 12*. Retrieved from http://www.ejhs.org/Volume12/Gender.htm

Fausto-Sterling, A. (2000). The five sexes, revisited. *The Sciences*, *40*(4), 18–23.

Fleming, M. Z., Jenkins, S. R., & Bugarin, C. (1980). Questioning current definitions of gender identity: Implications of the Bem Sex-Role Inventory for transsexuals. *Archives of Sexual Behavior*, *9*(1), 13–26.

Hansburry, G. (2005). The Middle Men: An Introduction to Transmasculine Identities. *Studies in Gender and Sexuality*, *6*(3), 241–264.

Hurtado, S., Alvarado, A. R., & Guillermo-Wann, C. (2015). Thinking about race: The salience of racial identity at two-and four-year colleges and the climate for diversity. *The Journal of Higher Education*, *86*(1), 127–155.

Hyde, J. S., Bigler, R. S., Tate, C. C., & Anders, S. M. van. (2018). The Future of Sex and Gender in Psychology: Five Challenges to the Gender Binary. *American Psychologist*. doi:http://dx.doi.org/10.1037/amp0000307

Jackson, L. A. (1985). Self-conceptions and gender role: The correspondence between gender-role categorization and open-ended self-descriptions. *Sex Roles*, *13*(9), 549–566.

Kane, J. M., & Mertz, J. E. (2002). Debunking myths about gender and mathematics performance. *Notices of the AMS, 59*, 10–21. doi:10.1090/noti790

Kuhn, M. H., & McPartland, T. S. (1954). An Empirical Investigation of Self-Attitudes. *American Sociological Review*, *19*(1), 68. doi:10.2307/2088175

Luhtanen, R., & Crocker, J. (1992). A Collective Self-Esteem Scale: Self-Evaluation of One's Social Identity. *Personality and Social Psychology Bulletin*, *18*(3), 302–318. doi:10.1177/0146167292183006

Moradi, B., & Parent, M. (2013). Assessment of Gender-Related Traits, Attitudes,Roles, Norms, Identity, and Experiences. In *APA handbook of testing and assessment in psychology, Vol. 2: Testing and assessment in clinical and counseling psychology* (pp. 463–485). APA.

Morawski, J. G. (1985). The measurement of masculinity and femininity: Engendering categorical realities. *Journal of Personality*, *53*(2), 196–223.

Nagoshi, J. L., Adams, K. A., Terrell, H. K., Hill, E. D., & Brzuzy, S. (2008). Gender differences in correlates of homophobia and transphobia. *Sex Roles*, *59*, 521–531. doi:10.1007/s11199-008-9458-7

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of social psychological attitudes, Vol. 1. Measures of personality and social psychological attitudes* (pp. 17–59). New York: Academic Press.

Penner, A. M., & CadwalladerOlsker, T. (2012). Gender differences in mathematics and science achievement across the distribution: What international variation can tell us about the role

of biology and society. In H. Forgasz & F. Rivera (Eds.), *Towards equity in mathematics education*.

Randel, A. (2002). Identity salience: A moderator of the relationship between group gender composition and work group conflict. *Journal of Organizational Behavior*, *23*(6), 749–766. doi:10.1002/job.163

Schellenberg, D., & Kaiser, A. (2018). The sex/gender distinction: Beyond F and M. In *APA handbook of the psychology of women* (Vol. 1, pp. 165–188). Washington, DC: APA.

Stewart, A. J., & McDermott, C. (2004). Gender in Psychology. *Annual Review of Psychology*, *55*, 519–44. doi:10.1146/annurev.psych.55.090902.141537

Strazds, A. (2015, August). The measurement of schematic breadth and salience of gender. Poster, Toronto, On., Canada.

Tate, C. C., Youssef, C. P., & Bettergarcia, J. N. (2014). Integrating the study of transgender spectrum and cisgender experiences of self-categorization from a personality perspective. *Review of General Psychology*, *18*(4), 302–312. doi:10.1037/gpr0000019

Thompson, V. L. S. (1999). Variables affecting racial-identity salience among African Americans. *The Journal of Social Psychology*, *139*(6), 748–761.

Unger, R. K. (1979). Toward a Redefinition of Sex and Gender. *American Psychologist*, *34*(11), 1085–1094.

Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, *140*(4), 1174–1204. doi:10.1037/a0036620

West, C., & Zimmerman, D. H. (1987). Doing Gender. *Gender & Society*, *1*(2), 125–151.

Westbrook, L., & Saperstein, A. (2015). New Categories Are Not Enough: Rethinking the Measurement of Sex and Gender in Social Surveys. *Gender & Society*, *29*(4), 534–560. doi:10.1177/0891243215584758

Wilcox, R. R. (2012). *Introduction to Robust Estimation and Hypothesis Testing* (3rd ed.). Retrieved from http://www.jstor.org/stable/2669876?origin=crossref

Figures

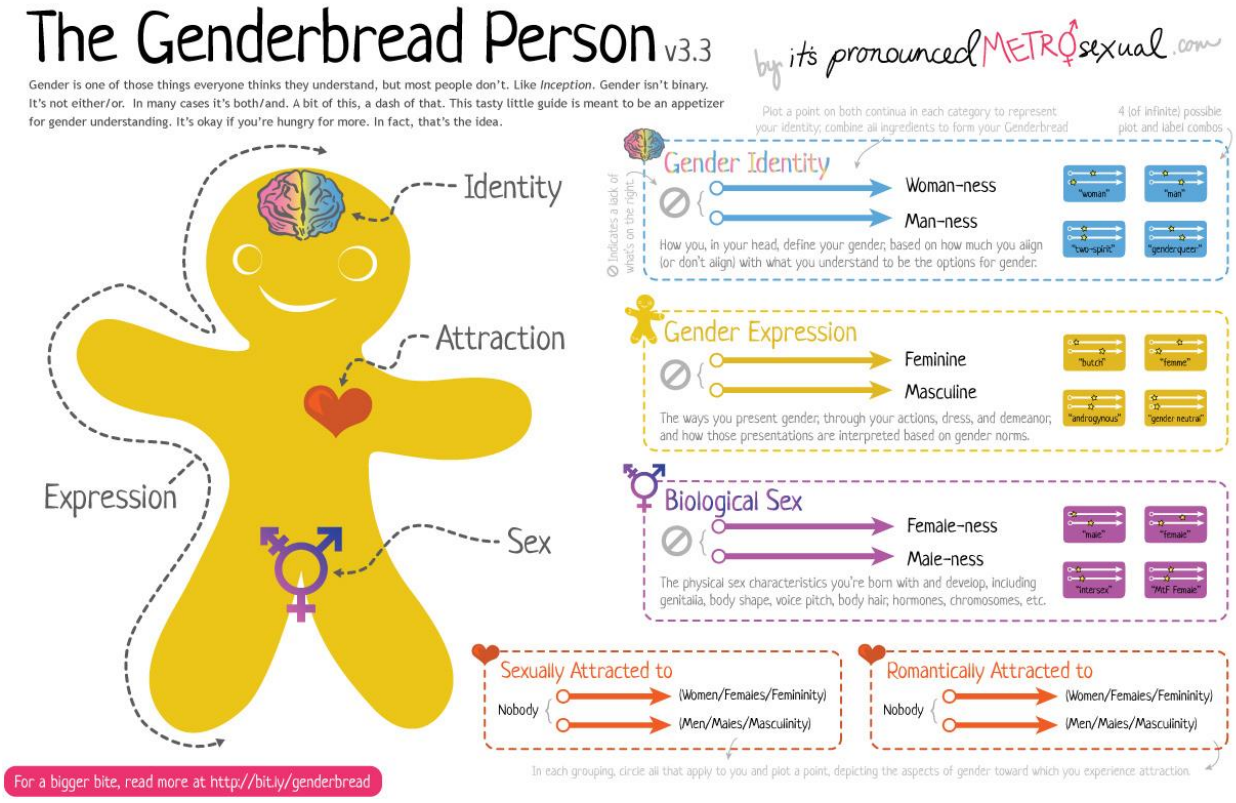

Figure 1 The Genderbread Person. The creators of this tool conceptualize gender as comprising two independent spectrums of masculinity and femininity. For a similar model more inclusive of non-binary individuals, see the Gender Unicorn at http://www.transstudent.org/gender/. Retrieved from http://itspronouncedmetrosexual.com/2015/03/the-genderbread-person-v3/ on June 22, 2015.

Appendix A

Tables for Full Gender x Schematic Breadth Model

Table 14

Schematic Breadth Poisson Multilevel Model Table, CM and Presentation as Reference

Predictor	β	Std. Error	z	p
Intercept	0.78	0.15	5.1	< .001
VF	0.02	0.01	3.93	< .001
CW vs CM	0.39	0.16	2.44	.015
NB vs CM	0.7	0.16	4.33	< .001
TM vs CM	0.28	0.15	1.82	.069
TW vs CM	0.36	0.15	2.35	.019
Role vs Pres	0.38	0.13	2.94	.003
Ident vs Pres	-0.11	0.15	-0.75	.451
CW vs CM x Role vs Pres	-0.14	0.15	-0.95	.344
NB vs CM x Role vs Pres	-0.87	0.16	-5.52	< .001
TM vs CM x Role vs Pres	-0.12	0.15	-0.84	.4
TW vs CM x Role vs Pres	-0.24	0.15	-1.66	.097
CW vs CM x Ident vs Pres	-0.37	0.17	-2.1	.036
NB vs CM x Ident vs Pres	-0.29	0.17	-1.68	.093
TM vs CM x Ident vs Pres	-0.25	0.17	-1.52	.128
TW vs CM x Ident vs Pres	-0.04	0.16	-0.26	.796

Table 15

Schematic Breadth Poisson Multilevel Model Table, CW and Presentation as Reference

Predictor	β	Std. Error	z	p
Intercept	1.17	0.12	10.06	< .001
VF	0.02	0.01	3.93	< .001
CM vs CW	-0.39	0.16	-2.44	.015
NB vs CW	0.31	0.13	2.49	.013
TM vs CW	-0.11	0.11	-0.98	.325
TW vs CW	-0.03	0.11	-0.26	.795
Role vs Pres	0.24	0.08	2.99	.003
Ident vs Pres	-0.48	0.09	-5.05	< .001
CM vs CW x Role vs Pres	0.14	0.15	0.95	.344
NB vs CW x Role vs Pres	-0.73	0.12	-6.03	< .001
TM vs CW x Role vs Pres	0.02	0.1	0.2	.845
TW vs CW x Role vs Pres	-0.1	0.1	-0.95	.341
CM vs CW x Ident vs Pres	0.37	0.17	2.1	.036
NB vs CW x Ident vs Pres	0.08	0.13	0.61	.542
TM vs CW x Ident vs Pres	0.11	0.12	0.92	.358
TW vs CW x Ident vs Pres	0.32	0.12	2.7	.007

Table 16

Schematic Breadth Poisson Multilevel Model Table, NB and Presentation as Reference

Predictor	*β*	Std. Error	z	*p*
Intercept	1.48	0.12	11.94	< .001
VF	0.02	0.01	3.93	< .001
CM vs NB	-0.7	0.16	-4.33	< .001
CW vs NB	-0.31	0.13	-2.49	.013
TM vs NB	-0.42	0.12	-3.63	< .001
TW vs NB	-0.34	0.12	-2.9	.004
Role vs Pres	-0.49	0.09	-5.37	< .001
Ident vs Pres	-0.4	0.09	-4.51	< .001
CM vs NB x Role vs Pres	0.87	0.16	5.52	< .001
CW vs NB x Role vs Pres	0.73	0.12	6.03	< .001
TM vs NB x Role vs Pres	0.75	0.11	6.58	< .001
TW vs NB x Role vs Pres	0.63	0.12	5.46	< .001
CM vs NB x Ident vs Pres	0.29	0.17	1.68	.093
CW vs NB x Ident vs Pres	-0.08	0.13	-0.61	.542
TM vs NB x Ident vs Pres	0.03	0.12	0.29	.773
TW vs NB x Ident vs Pres	0.24	0.12	2.13	.034

Table 17

Schematic Breadth Poisson Multilevel Model Table, TM and Presentation as Reference

Predictor	β	Std. Error	z	*p*
Intercept	1.05	0.11	9.68	< .001
VF	0.02	0.01	3.93	< .001
CM vs TM	-0.28	0.15	-1.82	.069
CW vs TM	0.11	0.11	0.98	.325
NB vs TM	0.42	0.12	3.63	< .001
TW vs TM	0.08	0.1	0.78	.435
Role vs Pres	0.26	0.07	3.79	< .001
Ident vs Pres	-0.36	0.08	-4.62	< .001
CM vs TM x Role vs Pres	0.12	0.15	0.84	.4
CW vs TM x Role vs Pres	-0.02	0.1	-0.2	.845
NB vs TM x Role vs Pres	-0.75	0.11	-6.58	< .001
TW vs TM x Role vs Pres	-0.12	0.1	-1.24	.214
CM vs TM x Ident vs Pres	0.25	0.17	1.52	.128
CW vs TM x Ident vs Pres	-0.11	0.12	-0.92	.358
NB vs TM x Ident vs Pres	-0.03	0.12	-0.29	.773
TW vs TM x Ident vs Pres	0.21	0.11	1.95	.051

Table 18

Schematic Breadth Poisson Multilevel Model Table, TW and Presentation as Reference

Predictor	β	Std. Error	z	*p*
Intercept	1.14	0.1	11.01	< .001
VF	0.02	0.01	3.93	< .001
CM vs TW	-0.36	0.15	-2.35	.019
CW vs TW	0.03	0.11	0.26	.795
NB vs TW	0.34	0.12	2.9	.004
TM vs TW	-0.08	0.1	-0.78	.435
Role vs Pres	0.14	0.07	1.95	.051
Ident vs Pres	-0.15	0.07	-2.07	.039
CM vs TW x Role vs Pres	0.24	0.15	1.66	.097
CW vs TW x Role vs Pres	0.1	0.1	0.95	.341
NB vs TW x Role vs Pres	-0.63	0.12	-5.46	< .001
TM vs TW x Role vs Pres	0.12	0.1	1.24	.214
CM vs TW x Ident vs Pres	0.04	0.16	0.26	.796
CW vs TW x Ident vs Pres	-0.32	0.12	-2.7	.007
NB vs TW x Ident vs Pres	-0.24	0.12	-2.13	.034
TM vs TW x Ident vs Pres	-0.21	0.11	-1.95	.051

Table 19

Schematic Breadth Poisson Multilevel Model Table, CM and Role as Reference

Predictor	β	Std. Error	z	*p*
Intercept	1.16	0.14	8.27	< .001
VF	0.02	0.01	3.93	< .001
CW vs CM	0.24	0.14	1.69	.091
NB vs CM	-0.17	0.16	-1.1	.27
TM vs CM	0.15	0.14	1.12	.264
TW vs CM	0.12	0.14	0.83	.407
Pres vs Role	-0.38	0.13	-2.94	.003
Ident vs Role	-0.49	0.14	-3.61	< .001
CW vs CM x Pres vs Role	0.14	0.15	0.95	.344
NB vs CM x Pres vs Role	0.87	0.16	5.52	< .001
TM vs CM x Pres vs Role	0.12	0.15	0.84	.4
TW vs CM x Pres vs Role	0.24	0.15	1.66	.097
CW vs CM x Ident vs Role	-0.22	0.16	-1.37	.171
NB vs CM x Ident vs Role	0.58	0.17	3.48	< .001
TM vs CM x Ident vs Role	-0.13	0.15	-0.85	.398
TW vs CM x Ident vs Role	0.2	0.15	1.3	.192

Table 20

Schematic Breadth Poisson Multilevel Model Table, CW and Role as Reference

Predictor	*β*	Std. Error	z	*p*
Intercept	1.4	0.11	12.33	< .001
VF	0.02	0.01	3.93	< .001
CM vs CW	-0.24	0.14	-1.69	.091
NB vs CW	-0.42	0.13	-3.18	.001
TM vs CW	-0.09	0.11	-0.84	.402
TW vs CW	-0.13	0.11	-1.18	.24
Pres vs Role	-0.24	0.08	-2.99	.003
Ident vs Role	-0.71	0.09	-7.79	< .001
CM vs CW x Pres vs Role	-0.14	0.15	-0.95	.344
NB vs CW x Pres vs Role	0.73	0.12	6.03	< .001
TM vs CW x Pres vs Role	-0.02	0.1	-0.2	.845
TW vs CW x Pres vs Role	0.1	0.1	0.95	.341
CM vs CW x Ident vs Role	0.22	0.16	1.37	.171
NB vs CW x Ident vs Role	0.81	0.13	5.99	< .001
TM vs CW x Ident vs Role	0.09	0.12	0.78	.434
TW vs CW x Ident vs Role	0.42	0.12	3.64	< .001

Table 21

Schematic Breadth Poisson Multilevel Model Table, NB and Role as Reference

Predictor	β	Std. Error	z	*p*
Intercept	0.98	0.13	7.48	< .001
VF	0.02	0.01	3.93	< .001
CM vs NB	0.17	0.16	1.1	.27
CW vs NB	0.42	0.13	3.18	.001
TM vs NB	0.33	0.12	2.65	.008
TW vs NB	0.29	0.12	2.3	.021
Pres vs Role	0.49	0.09	5.37	< .001
Ident vs Role	0.1	0.1	0.96	.335
CM vs NB x Pres vs Role	-0.87	0.16	-5.52	< .001
CW vs NB x Pres vs Role	-0.73	0.12	-6.03	< .001
TM vs NB x Pres vs Role	-0.75	0.11	-6.58	< .001
TW vs NB x Pres vs Role	-0.63	0.12	-5.46	< .001
CM vs NB x Ident vs Role	-0.58	0.17	-3.48	< .001
CW vs NB x Ident vs Role	-0.81	0.13	-5.99	< .001
TM vs NB x Ident vs Role	-0.72	0.12	-5.74	< .001
TW vs NB x Ident vs Role	-0.38	0.12	-3.13	.002

Table 22

Schematic Breadth Poisson Multilevel Model Table, TM and Role as Reference

Predictor	*β*	Std. Error	z	*p*
Intercept	1.31	0.11	12.35	< .001
VF	0.02	0.01	3.93	< .001
CM vs TM	-0.15	0.14	-1.12	.264
CW vs TM	0.09	0.11	0.84	.402
NB vs TM	-0.33	0.12	-2.65	.008
TW vs TM	-0.04	0.1	-0.38	.701
Pres vs Role	-0.26	0.07	-3.79	< .001
Ident vs Role	-0.62	0.08	-8.21	< .001
CM vs TM x Pres vs Role	-0.12	0.15	-0.84	.4
CW vs TM x Pres vs Role	0.02	0.1	0.2	.845
NB vs TM x Pres vs Role	0.75	0.11	6.58	< .001
TW vs TM x Pres vs Role	0.12	0.1	1.24	.214
CM vs TM x Ident vs Role	0.13	0.15	0.85	.398
CW vs TM x Ident vs Role	-0.09	0.12	-0.78	.434
NB vs TM x Ident vs Role	0.72	0.12	5.74	< .001
TW vs TM x Ident vs Role	0.33	0.1	3.17	.002

Table 23

Schematic Breadth Poisson Multilevel Model Table, TW and Role as Reference

Predictor	*β*	Std. Error	z	*p*
Intercept	1.27	0.1	12.49	< .001
VF	0.02	0.01	3.93	< .001
CM vs TW	-0.12	0.14	-0.83	.407
CW vs TW	0.13	0.11	1.18	.24
NB vs TW	-0.29	0.12	-2.3	.021
TM vs TW	0.04	0.1	0.38	.701
Pres vs Role	-0.14	0.07	-1.95	.051
Ident vs Role	-0.29	0.07	-3.99	< .001
CM vs TW x Pres vs Role	-0.24	0.15	-1.66	.097
CW vs TW x Pres vs Role	-0.1	0.1	-0.95	.341
NB vs TW x Pres vs Role	0.63	0.12	5.46	< .001
TM vs TW x Pres vs Role	-0.12	0.1	-1.24	.214
CM vs TW x Ident vs Role	-0.2	0.15	-1.3	.192
CW vs TW x Ident vs Role	-0.42	0.12	-3.64	< .001
NB vs TW x Ident vs Role	0.38	0.12	3.13	.002
TM vs TW x Ident vs Role	-0.33	0.1	-3.17	.002

Table 24

Schematic Breadth Poisson Multilevel Model Table, CM and Identity as Reference

Predictor	*β*	Std. Error	z	*p*
Intercept	0.67	0.16	4.22	< .001
VF	0.02	0.01	3.93	< .001
CW vs CM	0.02	0.17	0.12	.903
NB vs CM	0.41	0.17	2.41	.016
TM vs CM	0.02	0.16	0.14	.887
TW vs CM	0.32	0.16	1.98	.048
Role vs Ident	0.49	0.14	3.61	< .001
Pres vs Ident	0.11	0.15	0.75	.451
CW vs CM x Role vs Ident	0.22	0.16	1.37	.171
NB vs CM x Role vs Ident	-0.58	0.17	-3.48	< .001
TM vs CM x Role vs Ident	0.13	0.15	0.85	.398
TW vs CM x Role vs Ident	-0.2	0.15	-1.3	.192
CW vs CM x Pres vs Ident	0.37	0.17	2.1	.036
NB vs CM x Pres vs Ident	0.29	0.17	1.68	.093
TM vs CM x Pres vs Ident	0.25	0.17	1.52	.128
TW vs CM x Pres vs Ident	0.04	0.16	0.26	.796

Table 25

Schematic Breadth Poisson Multilevel Model Table, CW and Identity as Reference

Predictor	*β*	Std. Error	z	*p*
Intercept	0.69	0.13	5.48	< .001
VF	0.02	0.01	3.93	< .001
CM vs CW	-0.02	0.17	-0.12	.903
NB vs CW	0.39	0.14	2.81	.005
TM vs CW	0	0.13	0.02	.988
TW vs CW	0.29	0.12	2.36	.018
Role vs Ident	0.71	0.09	7.79	< .001
Pres vs Ident	0.48	0.09	5.05	< .001
CM vs CW x Role vs Ident	-0.22	0.16	-1.37	.171
NB vs CW x Role vs Ident	-0.81	0.13	-5.99	< .001
TM vs CW x Role vs Ident	-0.09	0.12	-0.78	.434
TW vs CW x Role vs Ident	-0.42	0.12	-3.64	< .001
CM vs CW x Pres vs Ident	-0.37	0.17	-2.1	.036
NB vs CW x Pres vs Ident	-0.08	0.13	-0.61	.542
TM vs CW x Pres vs Ident	-0.11	0.12	-0.92	.358
TW vs CW x Pres vs Ident	-0.32	0.12	-2.7	.007

Table 26

Schematic Breadth Poisson Multilevel Model Table, NB and Identity as Reference

Predictor	β	Std. Error	z	*p*
Intercept	1.08	0.13	8.37	< .001
VF	0.02	0.01	3.93	< .001
CM vs NB	-0.41	0.17	-2.41	.016
CW vs NB	-0.39	0.14	-2.81	.005
TM vs NB	-0.39	0.13	-3.07	.002
TW vs NB	-0.1	0.12	-0.77	.44
Role vs Ident	-0.1	0.1	-0.96	.335
Pres vs Ident	0.4	0.09	4.51	< .001
CM vs NB x Role vs Ident	0.58	0.17	3.48	< .001
CW vs NB x Role vs Ident	0.81	0.13	5.99	< .001
TM vs NB x Role vs Ident	0.72	0.12	5.74	< .001
TW vs NB x Role vs Ident	0.38	0.12	3.13	.002
CM vs NB x Pres vs Ident	-0.29	0.17	-1.68	.093
CW vs NB x Pres vs Ident	0.08	0.13	0.61	.542
TM vs NB x Pres vs Ident	-0.03	0.12	-0.29	.773
TW vs NB x Pres vs Ident	-0.24	0.12	-2.13	.034

Table 27

Schematic Breadth Poisson Multilevel Model Table, TM and Identity as Reference

Predictor	*β*	Std. Error	z	*p*
Intercept	0.69	0.11	6.06	< .001
VF	0.02	0.01	3.93	< .001
CM vs TM	-0.02	0.16	-0.14	.887
CW vs TM	0	0.13	-0.02	.988
NB vs TM	0.39	0.13	3.07	.002
TW vs TM	0.29	0.11	2.62	.009
Role vs Ident	0.62	0.08	8.21	< .001
Pres vs Ident	0.36	0.08	4.62	< .001
CM vs TM x Role vs Ident	-0.13	0.15	-0.85	.398
CW vs TM x Role vs Ident	0.09	0.12	0.78	.434
NB vs TM x Role vs Ident	-0.72	0.12	-5.74	< .001
TW vs TM x Role vs Ident	-0.33	0.1	-3.17	.002
CM vs TM x Pres vs Ident	-0.25	0.17	-1.52	.128
CW vs TM x Pres vs Ident	0.11	0.12	0.92	.358
NB vs TM x Pres vs Ident	0.03	0.12	0.29	.773
TW vs TM x Pres vs Ident	-0.21	0.11	-1.95	.051

Table 28

Schematic Breadth Poisson Multilevel Model Table, TW and Identity as Reference

Predictor	*β*	Std. Error	z	*p*
Intercept	0.98	0.11	9.35	< .001
VF	0.02	0.01	3.93	< .001
CM vs TW	-0.32	0.16	-1.98	.048
CW vs TW	-0.29	0.12	-2.36	.018
NB vs TW	0.1	0.12	0.77	.44
TM vs TW	-0.29	0.11	-2.62	.009
Role vs Ident	0.29	0.07	3.99	< .001
Pres vs Ident	0.15	0.07	2.07	.039
CM vs TW x Role vs Ident	0.2	0.15	1.3	.192
CW vs TW x Role vs Ident	0.42	0.12	3.64	< .001
NB vs TW x Role vs Ident	-0.38	0.12	-3.13	.002
TM vs TW x Role vs Ident	0.33	0.1	3.17	.002
CM vs TW x Pres vs Ident	-0.04	0.16	-0.26	.796
CW vs TW x Pres vs Ident	0.32	0.12	2.7	.007
NB vs TW x Pres vs Ident	0.24	0.12	2.12	.034
TM vs TW x Pres vs Ident	0.21	0.11	1.95	.051